



September 3, 2024

Hon. Ona T. Wang
Daniel Patrick Moynihan
United States Courthouse
500 Pearl St.
New York, NY 10007-1312

cc: All Counsel of Record Line (via ECF)

Re: New York Times Company v. Microsoft Corp., et al., Case No. 1:23-cv-11195-SHS-OTW

Dear Judge Wang:

Pursuant to Sections I(b) and II(b) of Your Honor's Individual Practices, Defendant OpenAI respectfully requests a pre-motion conference to address The New York Times's refusal to produce highly relevant discovery regarding its own adoption of, reliance on, and support of the very technology the Times claims to be illegal in this lawsuit: generative artificial intelligence.¹

Evidence regarding the Times's creation, use, and positions on generative AI generally is directly relevant to a number of issues in this case including, most prominently, OpenAI's fair use defense. Whether a use is fair under Section 107 of the Copyright Act turns, in part, on "the public benefits the copying will likely produce." *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 35 (2021); *see also Sega Enterprises Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1523 (9th Cir. 1992) (fair use analysis asks whether use "serves a public interest"). It is therefore not only relevant that the technology OpenAI introduced to the world "serves [the] public interest" generally, *see Accolade*, 977 F.2d at 1523, it is also relevant that the technology yields transformative and productive benefits for the enterprise of journalism specifically, *see Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 528 (2023) (use more likely to be fair if it has "further purpose or different character" than the allegedly copied work).

The Times will no doubt deny in this litigation the existence of such benefits—or at least minimize them. In fact, the Times has already done so by claiming that "GenAI" technology *as a whole* "[t]hreaten[s] [h]igh-[q]uality [j]ournalism." Dkt. 170 at 14. Outside of court, however, the Times has told a different story. Earlier this year, it hired a team to "leverag[e] generative [AI] . . . for the benefit of [its] journalists."² It also disclosed an intention to provide its journalists with "certain gen-A.I. tools"—without disclosing what tools it would provide or whether it planned to develop them in-house.³ The Times, in other words, has and will continue to strenuously deny the

¹ The parties conferred in good faith regarding the disputes addressed herein by telephone conference on May 14 and July 31, as well as by written correspondence both before and after those conferences. The parties' conferral efforts were unsuccessful. See Ex. E at 1, 5–6; Ex. F at 1.

² N.Y. Times, *Introducing the A.I. Initiatives Team* (May 9, 2024), <https://www.nytc.com/press/introducing-the-a-i-initiatives-team/>.

³ *Id.*; *see also* Sara Fischer, *Exclusive: NYT Plans to Debut New Generative AI Ad Tool Late This Year* (Feb. 20, 2024), <https://www.axios.com/2024/02/20/new-york-times-ai-ad-tool>.

public benefits of this technology in court—while simultaneously engaging in efforts to implement the very same technology in its newsroom and business operations.

Evidence of the Times’s efforts to create, deploy, or support the use of generative AI in its own newsroom and business operations is relevant not only to fair use, but also to the Times’s own claim that the mere existence of this technology is a threat to its business model and the enterprise of journalism. That evidence is also relevant to the Times’s assertion that it suffers “substantial harm” as a result of the deployment of this technology, *id.* ¶ 154—and, in turn, its claim for “billions of dollars in statutory and actual damages,” *id.* ¶ 9. For these reasons, OpenAI propounded discovery requests about “Generative AI”—defined to include both OpenAI’s services and other tools that rely on similar technology.⁴ OpenAI sought three categories of discovery about generative AI, regarding: (1) the Times’s use of generative AI (RFPs 44, 45, 46, 49); (2) the Times’s training of its own generative AI tools (RFPs 50, 51); and (3) the Times’s positions regarding generative AI (RFPs 4–6, 16, 24). *See* Ex. A at 7–8, 9, 11, 14–16.

When the Times made an analogous request for documents “concerning the impact of AI Models on Journalism . . . or any Generative AI Products . . . as a source of Journalism, substitute for Journalism, or product in competition with Journalism (including The Times’s Journalism)” — OpenAI did not attempt to limit this to just the Times’s journalism. *See* Ex. B at 7; Ex. C at 3.

Yet, the Times refused to produce any discovery relating to any tools other than those created by OpenAI on relevance grounds. *See* Ex. D at 3 (Times objection to the definition “Generative AI” and limiting responses to only OpenAI services). But the relevance of this evidence does not depend on whether the tools in question were built by OpenAI, the Times itself, or another company. It is relevant because it shows that the Times has acknowledged that the technology OpenAI pioneered and introduced to the world in late 2022 yields vast benefits for the public, for journalists, and for the Times itself. That, in turn, directly supports OpenAI’s fair use defense and undercuts the Times’s claim to have suffered a grave injury as a result of OpenAI’s innovations. The Times’s refusal to produce this discovery on relevance grounds is thus irreconcilable with Rule 26, which permits discovery into “any nonprivileged matter that is relevant to any party’s claim or defense and proportional to the needs of the case.” Fed. R. Civ. P. 26(b)(1) (emphasis added); *see also* *Ball v. Metro-N. Commuter R.R.*, No. 21-cv-6159, 2024 WL 1118783, at *2 (S.D.N.Y. Mar. 13, 2024) (“Relevance is construed broadly to encompass any matter that bears on, or that reasonably could lead to other matter that could bear on, any party’s claim or defense.” (cleaned up)).

The Times resists compliance with three categories of requests, all of which seek highly relevant evidence.

The Times’s Use of Generative AI (RFPs 44, 45, 46, 49). The Times has acknowledged the relevance of its own use of generative AI by agreeing to produce discovery relating to its use of OpenAI’s services. But, as noted above, the Times has refused to produce similar evidence regarding its use of generative AI services created by other companies. As discussed above, there

⁴ Ex. A at 2–3 (defining “Generative AI” to include “models, algorithms, or other systems that are trained using data to generate new content (including text, images, or other media), such as ChatGPT, Gemini, LLaMA, or Claude”).

is no basis for the Times to unduly narrow the scope of discoverable evidence in this manner. If the Times is using technology similar to OpenAI's, even if offered by another company, that would undermine the Times's assertion that tools like ChatGPT threaten journalism, just as evidence about the Times's use of ChatGPT would. Accordingly, the Times must produce that evidence.

The Times's Creation of its Own Generative AI (RFPs 50, 51). The Times has already suggested publicly that it has several species of generative AI tools in development, both for use in its newsroom and for use in its business operations.⁵ These reports suggest that the Times is in the process of leveraging the very technology it challenges in this lawsuit. That alone establishes that these efforts are highly relevant to OpenAI's defense, including its rebuttal of the Times's representation that generative AI "[t]hreaten[s] [h]igh-[q]uality [j]ournalism." Dkt. 170 at 14. Beyond that, OpenAI is entitled to know **(1)** the extent to which the Times has leveraged unlicensed data to create these tools (or pre-existing models trained on the same); **(2)** the extent to which the Times has admitted internally that there is no "workable market" for licensing the data needed to create an effective language model, *see Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 930 (2d Cir. 1994) (analyzing fourth fair use factor), and **(3)** the extent to which the Times has acknowledged that this technology could *enhance* the "traditional business models that supported quality journalism," Dkt. 170 ¶ 47. All of this evidence is directly relevant both to OpenAI's fair use defense and to the Times's damages claims. The Times must produce it.

The Times's Positions Regarding Generative AI (RFPs 4–6, 16, 24). The Times's positions expressed outside of litigation are relevant for similar reasons, as they have the potential to impeach the Times's assertions made to support its lawsuit. *See, e.g., Judd v. Take-Two Interactive Software, Inc.*, No. 07-cv-7932, 2009 WL 361958, at *1 (S.D.N.Y. Feb. 4, 2009) (ordering discovery that was "relevant as potential impeachment evidence"). That includes communications about "copyright and artificial intelligence" with third parties, as sought by RFP 24. Ex. A at 11. As explained *supra* p. 1, if such communications include concessions about generative AI broadly—*e.g.*, that generative AI technology is transformative or that it will yield significant public benefits—that would undermine the Times's contrary fair-use positions regarding OpenAI's generative AI services here.

Relevant discovery in this vein also includes the Times's knowledge about the training of third-party generative AI tools using the Times's works and the alleged reproduction of the Times's works by such tools, as sought by RFPs 4–6 and 16. Ex. A at 7–9. If the Times knew about multiple third parties using the Times's works to train generative AI tools but did nothing, that would suggest recognition by the Times of the reasons such training is protected by fair use—*e.g.*, that no workable market exists for licensing the volume of data required; that it offers sufficient public benefits; and that it stands to achieve purposes distinct from that of the underlying works. *See Texaco Inc.*, 60 F.3d at 930; *Oracle*, 593 U.S. at 35; *Warhol*, 598 U.S. at 528.

Accordingly, the Times should be compelled to provide discovery on the very subject of this lawsuit—generative AI—without narrowly limiting its production only to the materials that concern OpenAI. OpenAI is entitled to test whether the Times's statements and actions outside this lawsuit are consistent with the positions the Times is taking in this case.

⁵*See* N.Y. Times, *supra* note 2; Fischer, *supra* note 3.

Respectfully,

KEKER, VAN NEST &
PETERS LLP*

/s/ Michelle S. Ybarra
Michelle S. Ybarra

LATHAM & WATKINS
LLP

/s/ Elana Nightingale Dawson
Elana Nightingale Dawson

MORRISON &
FOERSTER LLP

/s/ Vera Ranieri
Vera Ranieri

* All parties whose electronic signatures are included herein have consented to the filing of this document.